

# Uneven Distribution of Mutational Variance Across the Transcriptome of *Drosophila serrata* Revealed by High-Dimensional Analysis of Gene Expression

Emma Hine,<sup>\*,1</sup> Daniel E. Runcie,<sup>†</sup> Katrina McGuigan,<sup>\*</sup> and Mark W. Blows<sup>\*</sup>

<sup>\*</sup>School of Biological Sciences, The University of Queensland, St Lucia, Queensland 4068, Australia and <sup>†</sup>Department of Plant Sciences, University of California, Davis, California

ORCID IDs: 0000-0002-2257-4430 (E.H.); 0000-0002-3008-9312 (D.E.R.); 0000-0002-0525-2202 (K.M.); 0000-0002-1065-5524 (M.W.B.)

**ABSTRACT** There are essentially an infinite number of traits that could be measured on any organism, and almost all individual traits display genetic variation, yet substantial genetic variance in a large number of independent traits is not plausible under basic models of selection and mutation. One mechanism that may be invoked to explain the observed levels of genetic variance in individual traits is that pleiotropy results in fewer dimensions of phenotypic space with substantial genetic variance. Multivariate genetic analyses of small sets of functionally related traits have shown that standing genetic variance is often concentrated in relatively few dimensions. It is unknown if a similar concentration of genetic variance occurs at a phenome-wide scale when many traits of disparate function are considered, or if the genetic variance generated by new mutations is also unevenly distributed across phenotypic space. Here, we used a Bayesian sparse factor model to characterize the distribution of mutational variance of 3385 gene expression traits of *Drosophila serrata* after 27 generations of mutation accumulation, and found that 46% of the estimated mutational variance was concentrated in just 21 dimensions with significant mutational heritability. We show that the extent of concentration of mutational variance into such a small subspace has the potential to substantially bias the response to selection of these traits.

**KEYWORDS** pleiotropy; gene expression; mutation; genetic constraints

THE phenotype of an organism can essentially be measured in an infinite number of ways (Houle 2010), and the number of nucleotide changes that could occur to alter the phenotype exceeds  $10^6$  for organisms as complex as *Drosophila* (Johnson and Barton 2005). There are two reasons to suspect that for any large number of measured traits,  $p$ , only  $n \ll p$  independent (orthogonal) dimensions will exhibit levels of genetic variance comparable to the substantial levels commonly measured in individual traits. First, costs associated with organismal complexity suggest an evolutionary advantage to limiting  $n$  (Orr 2000); the probability of new mutations being favorable, the magnitude of the favorable effects, and the chances of fixation all decrease with increasing  $n$ . The disadvantages of high  $n$  are predicted to increase

almost exponentially as  $n$  exceeds more than a few hundred, leading to a rapidly declining rate of adaptation in more complex organisms (Orr 2000). Second, for a population close to its optimum, and evolving under Gaussian stabilizing selection acting independently on each trait, genetic load will reduce mean fitness as the number of independently selected traits increases (Barton 1990; Johnson and Barton 2005). Based on published estimates of heritability and stabilizing selection, this basic model of mutation and selection predicts that population fitness will decrease to unsustainable levels when  $n$  exceeds a few hundred (Barton 1990; Johnson and Barton 2005).

If  $n \ll p$ , as implied by Fisher's geometric model (Fisher 1930) and theoretical models of selection and quantitative variation, much of the genetic variance will be confined to the smaller  $n$ -dimensional subspace. The distribution of genetic variance is characterized by the empirical spectral distribution of the  $p \times p$  genetic covariance matrix ( $\mathbf{G}$ ). When  $n$  is much less than  $p$ , the first  $n$  eigenvalues of  $\mathbf{G}$  will account for a greater proportion of the genetic variance than expected when all dimensions of phenotypic space contain substantial

Copyright © 2018 by the Genetics Society of America

doi: <https://doi.org/10.1534/genetics.118.300757>

Manuscript received January 25, 2018; accepted for publication May 31, 2018; published Early Online June 8, 2018.

Supplemental material available at Figshare: <https://doi.org/10.25386/genetics.6460916>.

<sup>1</sup>Corresponding author: School of Biological Sciences, The University of Queensland, St Lucia, Queensland 4068, Australia. E-mail: [e.hine@uq.edu.au](mailto:e.hine@uq.edu.au)

genetic variance. For a set of  $p$  traits, strengthening the covariance relationships among them increases the genetic variance contained in the  $n$ -dimensional subspace, in turn decreasing the genetic variance contained in the complementary  $(p-n)$ -dimensional subspace. This uneven distribution of genetic variance will affect the predicted response to selection; as  $n$  decreases relative to  $p$ , the evolutionary response is increasingly biased toward those  $n$  directions with greater genetic variance (Walsh and Blows 2009).

There is now considerable evidence from many empirical studies of small sets of functionally related traits that a substantial proportion of genetic variance is restricted to a smaller subspace of the measured phenotype (Blows and McGuigan 2015). Although this observation is generally consistent across different trait types and organisms, there are some important limitations to our current understanding of the distribution of genetic variance. Sets of functionally related traits, which are most often analyzed in multivariate quantitative genetic studies, might be expected to exhibit greater pleiotropic covariance than typical among the disparate function traits that comprise an organism. These studies may therefore give a misleading impression of the extent of the concentration of genetic variance in a smaller subspace. It is also unclear how the concentration of genetic variance may scale to a phenome-wide level, as high-dimensional genetic analyses are few (Runcie and Mukherjee 2013; Blows *et al.* 2015).

We know much less about the extent of pleiotropy generated by new mutations, and specifically whether mutational variance, the ultimate source of genetic variance, is concentrated in a small subspace or more evenly distributed than standing genetic variance. Under Fisher's model, a mutation possibly affects all traits to some extent (universal pleiotropy), but some traits are affected much more than others (Orr 2000; Martin and Lenormand 2006), leading to pleiotropic covariance among traits and, consequently, a concentration of mutational covariance in a smaller subspace. While only a few studies have directly characterized the distribution of eigenvalues of the mutational covariance matrix (**M**), each has reported an uneven distribution of mutational variance, with a small number of eigenvalues accounting for the majority of the mutational variance (Camara and Pigliucci 1999; Latimer *et al.* 2014; McGuigan *et al.* 2014). In the only empirical study we are aware of to determine the relative size of the  $n$ -dimensional subspace of **M**, Houle and Fierst (2013) showed that for  $p = 20$  wing shape traits of *Drosophila melanogaster*,  $n$  ranged between 7 and 19, suggesting  $n < p$  as a consequence of pleiotropy. How these results relate to a wider set of traits, and if the distribution of mutational variance at a phenome-wide level is similarly uneven, remains to be determined.

Here, we utilize a recently developed analytical framework to investigate the distribution of mutational variance across a very large number of gene expression traits with a wide range of putative functions. As the number of parameters required to estimate covariance matrices scales with  $p$  by  $p \times (p + 1)/2$ ,

estimation of such matrices within the context of quantitative genetic experimental designs is particularly challenging (Blows *et al.* 2015). The Bayesian sparse factor (BSF) model of Runcie and Mukherjee (2013) is an extension of factor analytic modeling that overcomes the estimation challenge via two key sparsity assumptions: first, few latent factors are assumed to underlie the phenotypic variance, and second, most latent factors are assumed to affect relatively few individual traits (Runcie and Mukherjee 2013).

We apply the BSF model to investigate mutational covariance among 3385 gene expression traits after 27 generations of mutation accumulation (MA) in 41 lines of *D. serrata* (McGuigan *et al.* 2014). Although some of the expressed genes are involved in the same biological processes (Blows *et al.* 2015; Collet *et al.* 2018), there is no *a priori* reason to expect strong covariance among this functionally diverse set of traits. In a previous analysis of these data, frequent mutational covariance was established among small  $p = 5$  random sets of these 3385 expression phenotypes, indicating that  $n \ll p$  to some unknown extent (McGuigan *et al.* 2014). The BSF model allowed us to directly investigate how many expression traits are affected by a common latent factor (the transcriptome-wide extent of mutational covariance), and whether the total observed mutational variance in this experiment is consistent with many latent factors (mutations) affecting each trait independently or with few common factors affecting many traits simultaneously. We then investigated the potential evolutionary consequences of the extent of mutational covariance by determining the predicted bias in the response to selection that would be generated by the observed distribution of mutational variance.

## Materials and Methods

The data and the experimental design from which it was obtained have been described elsewhere (McGuigan *et al.* 2014). Briefly, 100 lines were set up from a single inbred line and maintained for 27 generations of an MA experiment. In each line, one son and one daughter of a single brother-sister pair became the parents of the next generation. In the 28th generation, 11,604 gene expression traits were measured (using five different probes per trait, with each probe duplicated) on two biological replicates for each of the 41 surviving MA lines. McGuigan *et al.* (2014) conducted univariate mixed effects models on the 11,604 gene expression traits and reported that 3385 had nonzero estimates of among-line (mutational) variances, with 1035 significant at  $\alpha = 0.05$ , and 533 remaining significant after false discovery rate (FDR) correction. Of 677 randomly allocated five-trait sets, a common component of among-line variance was significant for 245 (145 at FDR).

## Statistical analysis

A recent advance in the analysis of high-dimensional genetic data has made it possible to estimate the mutational variances of, and covariance among, all 3385 traits in a single analysis.

The BSF genetic (BSFG) model (Runcie and Mukherjee 2013) uses factor analysis to partition the total phenotypic variance into common phenotypic variance arising from latent factors and trait-specific phenotypic variance. At the same time, the BSF model utilizes information from the experimental design to further partition these variances into their genetic and residual (environmental, technical *etc.*) components. Two key assumptions distinguish the BSF from traditional factor analysis and make it possible to include so many traits in the one analysis with relatively small sample sizes. First, few latent factors are assumed to underlie the phenotypic variance. Second, most latent factors are assumed to affect relatively few individual traits (Runcie and Mukherjee 2013).

We refer the reader to Runcie and Mukherjee (2013) for a detailed description of the BSF genetic model for a one-way experimental design utilizing pedigree information in an animal model. Here, we modify this model to accommodate the experimental design outlined above. Note that Runcie and Mukherjee (2013) use the BSF model to infer the genetic variance–covariance matrix ( $\mathbf{G}$ ), whereas here we are estimating mutational (co)variances, which can be summarized with the covariance matrix  $\mathbf{M}$ . Further, Runcie and Mukherjee’s model assumed one probe per gene, and so we extended the model here to allow for multiple uncorrelated probes per gene.

Prior to analysis, we took the  $\log_{10}$  of the mean of two measurements per probe on an array. A genetic variant was known to be segregating in this set of MA lines (McGuigan *et al.* 2014), so we extracted the residuals from simple linear models for each set of  $\log_{10}$  mean expression measures, and standardized them to unit variance. Let  $y_{defg}$  represent these standardized residuals for line  $d \in 1, 2, \dots, D = 41$ , replicate  $e \in 1, 2$ , gene  $f \in 1, 2, \dots, F = 3385$ , and probe  $g \in 1, 2, \dots, G = 5$ .  $y_{defg}$  is modeled as

$$y_{defg} = \mu_f + \text{Probe}_{fg} + \text{Line}_{df} + \text{Rep}_{def} + \epsilon_{defg} \quad (1)$$

where  $\mu_f$  is the global mean for gene  $f$ ,  $\text{Probe}_{fg}$  is the mean effect of the two replicates per array of probe  $g$  within gene  $f$ ,  $\text{Line}_{df}$  is the mean effect of line  $d$  on gene  $f$ ,  $\text{Rep}_{def}$  is the mean effect for replicate  $e$  on gene  $f$ , and  $\epsilon_{defg}$  is the residual error. The line, replicate, and residual effects are considered random (with modeled variance), and the mean and probe effects are considered fixed.

To model the mutational covariance among multiple gene expression traits, we assume that the random effects have the following distributions:

$$\begin{aligned} \text{Line}_d &\sim \text{MVN}_F(0, \mathbf{M}) \\ \text{Rep}_{de} &\sim \text{MVN}_F(0, \mathbf{R}) \\ \epsilon_{defg} &\sim \text{N}(0, \sigma_{fg}^2) \end{aligned}$$

where the period in the index represents a vector formed from all  $F = 3385$  genes,  $\text{MVN}_F$  is the  $F$ -dimensional multivariate normal density, and  $\mathbf{M}$  and  $\mathbf{R}$  are covariance matrices of the gene expression traits caused by mutation and by microarray

technical errors plus biological noise, respectively. Mutation and residuals are assumed uncorrelated across lines and replicates, respectively.

Equation 1 can be reparameterized using hierarchical centering. This technique reduces correlations among parameters making Markov Chain Monte Carlo (MCMC) sampling more efficient. In particular, we introduce the latent variable  $w_{def}$  to model the gene-level effects in each rep:

$$\begin{aligned} y_{defg} &= w_{def} + \text{Probe}_{fg} + \epsilon_{defg} \\ w_{def} &= \mu_f + \text{Line}_{df} + \text{Rep}_{def} \end{aligned} \quad (2)$$

In Equation 2,  $\epsilon_{defg}$  are residuals of the probe measurements around a “true” transcript expression, and  $\text{Rep}_{def}$  are residuals of the transcript expression around the line means. Parameterized as such, the model for  $w_{def}$  has an identical form to the model presented in Runcie and Mukherjee (2013). Therefore, we embedded the BSFG Gibbs sampler from Runcie and Mukherjee (2013) in a Gibbs sampler for the  $\text{Probe}_{fg}$  and  $\sigma_{fg}^2$  parameters, sampling each parameter individually conditionally on the current sample of  $w_{def}$ .

As in a traditional factor analysis, we assume that any phenotypic covariance among the observed traits is caused by  $k$  latent factors, which we refer to as phenotypic common factors (PCFs), and allow specific mutational and residual variances for individual traits such that the phenotypic covariance matrix can be expressed as:

$$\mathbf{P} = \mathbf{\Lambda}\mathbf{\Lambda}^T + \psi_m + \psi_r$$

where  $\mathbf{\Lambda}$  is the  $F \times k$  matrix of PCF loadings, and  $\psi_m$  and  $\psi_r$  are the diagonal matrices of specific mutational and residual variances, respectively. Under the BSF model, the mutational covariance is modeled by allowing each latent PCF to have its own heritability. This heritability is the cumulative effect of new mutations, and it should not be confused with the per generation mutational heritability often reported in MA studies,  $V_M/V_E$ , where  $V_M$  is the among-line variance divided by twice the number of generations of MA.

The mutational ( $\mathbf{M}$ ) covariance matrix can then be recovered through the equation:

$$\mathbf{M} = \mathbf{\Lambda}\mathbf{\Sigma}_{h^2}\mathbf{\Lambda}^T + \psi_m \quad (3)$$

where  $\mathbf{\Sigma}_{h^2}$  is the  $k \times k$  diagonal matrix of PCF heritabilities. In our results, we refer to the common, specific, and total mutational variances of the individual traits, which correspond to the diagonals of  $\mathbf{\Lambda}\mathbf{\Sigma}_{h^2}\mathbf{\Lambda}^T$ ,  $\psi_m$ , and  $\mathbf{M}$ , respectively. Further, we present mutational variances of the estimated PCFs, rather than attempting to interpret a spectral decomposition of  $\mathbf{M}$ .

**Parameter expansion:** When assessing model convergence (see below), we noticed poor mixing of  $\mathbf{\Lambda} = [\lambda_{ij}]$ , and strongly negative posterior correlations between individual  $\lambda_{ij}$  and the SD of the individual PCF scores ( $f_j$ ). To combat this, we implemented the parameter expansion for factor models proposed

by Ghosh and Dunson (2009). We introduced a new parameter  $\Psi_F$ , a diagonal matrix, and replaced the parameters  $\Lambda$ ,  $F$ , and  $F_a$  in the original BSFG model (Runcie and Mukherjee 2013) with working parameters  $\Lambda^*$ ,  $F^*$ , and  $F_a^*$  with the following distributions:

$$\begin{aligned}\Lambda^* &= [\lambda_{ij}^*], \quad \lambda_{ij} \sim N(0, \phi_{ij}^{-1} \tau_j^{-1}) \\ F^* &\sim MN(ZF_a^*, I, \Psi_F^{1/2} (I - \Sigma_{h^2}) \Psi_F^{1/2}) \\ F_a^* &\sim MN(0, A, \Psi_F^{1/2} \Sigma_{h^2} \Psi_F^{1/2}) \\ \Psi_F &= \text{Diag}(\psi_{F_j}), \quad \psi_{F_j} \sim \text{Ga}(a_F, b_F)\end{aligned}$$

where  $\phi_{ij}$ ,  $\tau_j$ ,  $Z$ ,  $A$ , and  $\Sigma_{h^2}$  are defined as in Runcie and Mukherjee (2013).

For each posterior sample, we rescale  $F^*$ ,  $F_a^*$ , and  $\Lambda^*$  to recover the original parameters and to ensure that the variance of each ( $f_j$ ) is equal to one:

$$\begin{aligned}\Lambda^{(i)} &= \Lambda^{*(i)} \Psi_F^{(i)1/2} \hat{\Sigma}_F^{(i)1/2} \\ F^{(i)} &= F^{*(i)} \Psi_F^{(i)-1/2} \hat{\Sigma}_F^{(i)-1/2} \\ F_a^{(i)} &= F_a^{*(i)} \Psi_F^{(i)-1/2} \hat{\Sigma}_F^{(i)-1/2}\end{aligned}$$

where  $\hat{\Sigma}_F$  is a diagonal matrix of the empirical variance of  $F$  from each posterior sample.

Further investigations of Gibbs sampler performance identified the  $\delta_j, j \in 1, \dots, k$  parameters as exhibiting especially poor mixing and strong correlations among parameters. Since updates of individual  $\delta_j$  parameters are relatively fast, we iterated through the sampling of these parameters  $100 \times$  per iteration of the remaining parameters in the Gibbs sampler.

**Prior distributions and hyperparameters:** We used the same prior distributions as in Runcie and Mukherjee (2013) (with some changes to the hyperparameters, discussed below), and refer the reader to their paper for a detailed description. A summary of the prior distributions outlined below, and the values we ultimately used for their hyperparameters, are listed in Supplemental Material, Table S1. Briefly, the PCFs are modeled through a hierarchical distribution on the PCF loadings. Each  $\lambda_{ij}$  is assigned a normal distribution as a prior with a unique variance drawn from an inverse  $\gamma$  distribution. Modeling the PCF loadings in this way imposes sparsity on the PCFs and ensures that the variance explained by successive PCFs decreases to zero at some sufficiently large  $k^*$ .

The remaining parameters are estimated with more straightforward prior distributions. The heritability of the PCFs are modeled with a 50% probability of  $h_j^2 = 0$  and the remaining 50% distributed equally across a discrete set of evenly spaced values in the interval  $0 < x \leq 1$ . The mutational and residual specific variances, forming the diagonal elements of  $\psi_m$  and  $\psi_r$ , are modeled with inverse  $\gamma$  prior distributions. We also used inverse  $\gamma$  prior distributions to model the residual probe variances that form the diagonal

of  $\Sigma_y$ , which were not present in the original BSF genetic model presented in Runcie and Mukherjee (2013).

We explored the effect of changing the hyperparameter values used in Runcie and Mukherjee (2013) for three aspects of the prior distribution relating to the PCF trait loadings (Table S2). We ran the BSF model with five different random seeds for each of nine sets of values for the three hyperparameters. For each analysis, we retained 1000 samples with a thinning rate of 100 after a burn-in period of 300,000 samples. The total mutational variance estimated for the 3385 gene expression traits was consistent among the 45 analyses, with correlations between pairs of analyses ranging from 0.95 to 1.00. In contrast, the number of heritable PCFs and the amount of mutational variance accounted for by PCFs (common mutational variance) was sensitive to the hyperparameter values (Table S2).

**Model convergence:** From the 45 prior analyses, we identified the one that detected the greatest number of heritable PCFs (Table S2) and continued the chain. After discarding an additional 100,000 samples (taking the burn-in period to 400,000), then retaining 1000 samples at a thinning rate of 100, we were satisfied with the convergence diagnostics of the model. Autocorrelation statistics for the PCF heritabilities and specific variances are summarized in Figure S1. A single specific mutational variance exceeded our nominal autocorrelation threshold of 0.2. Autocorrelations for 630 of 152,325 PCF trait loadings exceeded 0.2 (Figure S2). We concluded from trace plots of these parameters (specific mutational variance not shown; samples of PCF trait loadings presented in Figure S3) that the high autocorrelations were unlikely to be cause for concern. We note that the PCFs most represented among the highly autocorrelated trait loadings (PCFs 1, 2, 4, 5, 7, 8, 17, and 18; Figure S2) are among those with very low (median posterior sample = 0), nonsignificant estimates of PCF heritability, which we do not interpret.

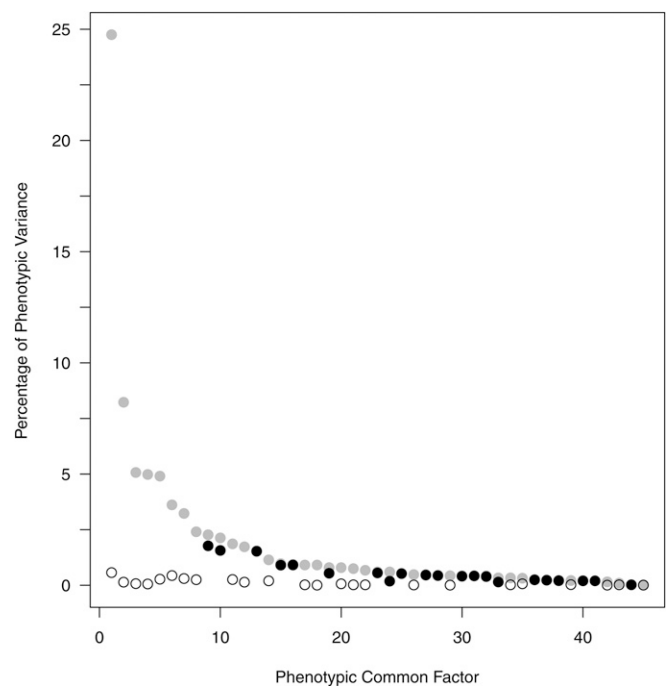
**Significance testing:** We use a local false sign rate (LFSR) approach to declare the significance of PCF heritabilities and trait loadings. LFSR is the probability of assigning the incorrect sign to an estimate. For PCF heritabilities, we assign LFSR as the proportion of posterior samples equal to zero. For PCF trait loadings, we assign LFSR values as the proportion of posterior samples equal to zero, or on the other side of zero from the median posterior sample. For example, if we declared all trait loadings with 75% of the posterior samples  $> 0$  to be positive, we expected to be wrong 25% of the time, so  $LFSR = 0.25$ . We report the number of trait loadings per significantly heritable PCF with  $LFSR < 0.025$  (being two-tailed tests) in Table S3. We also compute an average error rate for each set of heritable PCF trait loadings (Stephens 2017), and report the number of trait loadings for which the resulting  $s$ -value remains  $< 0.005$  (Table S3). The  $s$ -value is analogous to Storey's  $q$ -value (Storey 2003; Stephens 2017), and implies that we can expect 99% of the trait loadings declared significant to actually be different from zero.

**Randomizations:** We conducted two types of randomizations, referred to as type 1 and type 2. For type 1 randomizations, we tested whether the BSF model would allocate mutational variance to PCFs in the absence of a true covariance signal. To do this, we preserved the mutational variance of individual traits but disrupted trait covariances by shuffling line replicate pairs within traits. That is, we reassigned probe measurement  $y_{defg}$  to  $y_{d^*efg}$ , where all five probe measurements of the  $f$ th gene for both replicates of the  $d$ th line ( $d = 1, 2, \dots, 41$ ) take on the corresponding measurements for the  $f$ th gene of the  $d^*$ th line, where  $d^*$  is the  $d$ th element of the shuffled vector of integers 1–41, shuffled independently for each gene expression trait. For type 2 randomizations, we tested whether the BSF model would spuriously allocate heritability to PCFs given the observed trait covariances, but in the absence of mutational variance. We preserved the covariance structure among traits but disrupted the mutational variance of individual traits by shuffling the rows (representing individual replicates) of the  $82 \times 16,925$  data matrix. For each type of randomization, we generated 100 data sets. We ran the BSF model for each data set, collecting 1000 samples at a thinning rate of 100 after 300,000 burn-in samples.

**Predicted response to selection:** To explore how the distribution of the total mutational variance (sum of both common and specific mutational variances as described in Equation 3) might affect the adaptive potential of this population, we used  $\mathbf{M}$  in place of  $\mathbf{G}$  in the multivariate breeders' equation (BE),  $\Delta\mathbf{z} = \mathbf{G}\beta$ . We calculated the predicted response to selection,  $\Delta\mathbf{z}$ , for 1,000,000 random selection gradients,  $\beta$ , and subsequently determined the angle between  $\beta$  and  $\Delta\mathbf{z}$  as a measure of evolutionary bias, ranging from zero (no bias) to  $90^\circ$  (strong bias). Each of the 3385 elements of the million random  $\beta$ s was drawn from a uniform distribution between  $-1$  and  $1$ . For each of 1000 posterior samples of  $\mathbf{M}$ , we calculated  $\Delta\mathbf{z}$  for a subset of 1000 random  $\beta$ s, generating the total distribution of 1,000,000 predicted evolutionary responses to random directions of selection. We compared the distribution of angles between  $\beta$  and  $\Delta\mathbf{z}$  to the distribution generated when  $\beta$  was applied to  $\mathbf{M}$  with a null covariance structure (*i.e.*, our first randomization as outlined above). For the randomized data, we applied the same set of 1,000,000  $\beta$ s, allocating 100  $\beta$ s to each of 100 posterior samples of  $\mathbf{M}$  from each of the 100 randomized data sets.

### Data availability

The original data set is available at the Gene Expression Omnibus under accession number GSE49815. The subset of gene expression traits used in this analysis is available at [espace.library.uq.edu.au](http://espace.library.uq.edu.au) under accession number uql.2017.783. MATLAB code to implement the BSF analysis, MATLAB output, R code used to process the output, and univariate REML estimates are also available at [espace.library.uq.edu.au](http://espace.library.uq.edu.au) under accession numbers uql.2018.121, uql.2018.111, uql.2018.101, and uql.2018.113. Supplemental



**Figure 1** Phenotypic and mutational variances of the phenotypic common factors (PCFs), as a proportion of the total phenotypic variance across the 3385 gene expression traits. Before dividing by the total phenotypic variance, the phenotypic variance (gray filled circles) of a given PCF is defined as the sum of its squared trait loadings. The phenotypic variance is multiplied by the corresponding PCF heritability to give the mutational variance attributable to the PCF. Black filled and unfilled circles depict mutational variances for PCFs with significant ( $p < 0.01$ ) and nonsignificant PCF heritabilities, respectively.

material available at Figshare: <https://doi.org/10.25386/genetics.6460916>.

## Results and Discussion

After detailed investigation of the effect of several prior distribution hyperparameter values, and confirmation of convergence of the selected final BSF model (*Materials and Methods*), we conducted two sanity checks to ensure the final BSF model provided a good representation of the known characteristics of the experimental design and data. First, we used the Marcenko–Pastur law from random matrix theory for large sample covariance matrices to form an expectation of how many PCFs we could reasonably expect the BSF model to detect (Nadakuditi and Edelman 2008). Our experimental design of 82 independent experimental units and 3385 traits should have sufficient power to detect, with high probability, a minimum of 37 dimensions of phenotypic variance (Figure S4). The BSF model identified 45 PCFs (Figure 1), indicating that the model had converged on a realistic number of PCFs given our experiment. Second, the mutational variances for the 3385 individual expression traits previously estimated from standard univariate REML analyses (McGuigan *et al.* 2014) were highly correlated ( $r = 0.84$ , d.f. = 3383,  $P < 0.001$ ) with the BSF estimates (Figure S5), indicating that

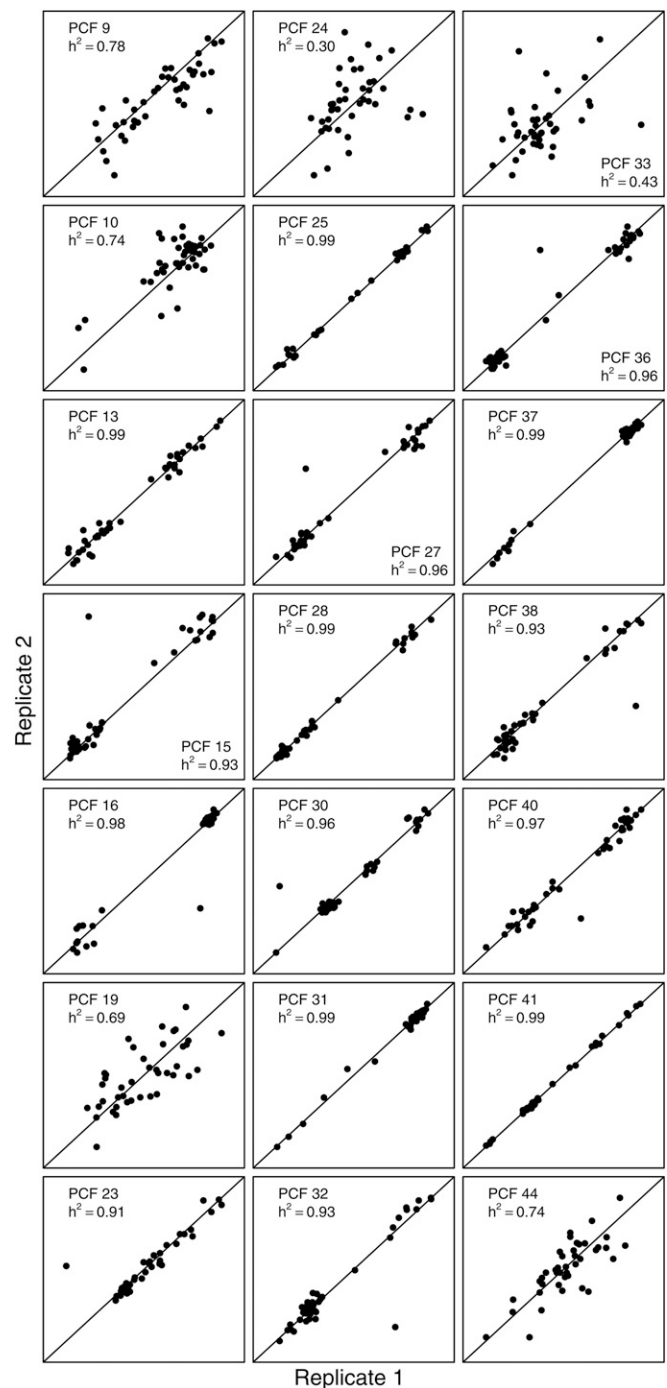
the mutational variance in individual expression traits was recovered well by the more complex BSF model.

Mutational heritability was significant at  $s < 0.01$  for 21 of the 45 PCFs (Figure 1 and Figure S6), with  $0.30 \leq h^2 \leq 0.99$  (Figure 2). We compared the observed PCF heritabilities and their LFSR values to those obtained for the randomized data sets (Figure 3). For the type 1 randomizations, which simulate traits with uncorrelated mutational effects, a single PCF from 3 of the 100 data sets was, by chance, highly heritable ( $>0.8$ , Figure 3A). The remaining PCF heritabilities estimated from the type 1 randomized data sets were all  $<0.1$ , well below the smallest observed PCF heritability. For the type 2 randomizations, which simulate traits with phenotypic covariance but no mutational variance, most of the PCF heritabilities were concentrated near zero. Only 1.4% of the type 2 randomized data PCF heritabilities exceeded the smallest observed PCF heritability. These results demonstrate that, despite the limited genetic d.f., there is only a small probability that any of the observed heritable PCFs were spuriously generated by the BSF model.

There was considerable variation in the way individual traits contributed to the heritable PCFs. Overall, 1193 of the 3385 expression traits contributed significantly at  $s < 0.005$  to at least one of the 21 heritable PCFs (Figure 4). The number of trait loadings on a given PCF ranged from 3 to 266 ( $s < .005$ ; Figure 4 and Table S3). Most expression traits contributed significantly to only a single PCF, but some contributed to up to four (Table S3). In general, the extent of shared expression traits across PCFs scaled with the size of the PCF; the three largest PCFs had the most individual expression traits that contributed to multiple PCFs (Table S3). Furthermore, the two heritable PCFs associated with the greatest phenotypic and mutational variance (PCFs 9 and 10) displayed individual trait loadings that were biased to one side of zero. This is consistent with the major axis of standing genetic variance in gene expression in this species (Blows *et al.* 2015), which is also characterized by a majority of trait loadings being biased in one direction.

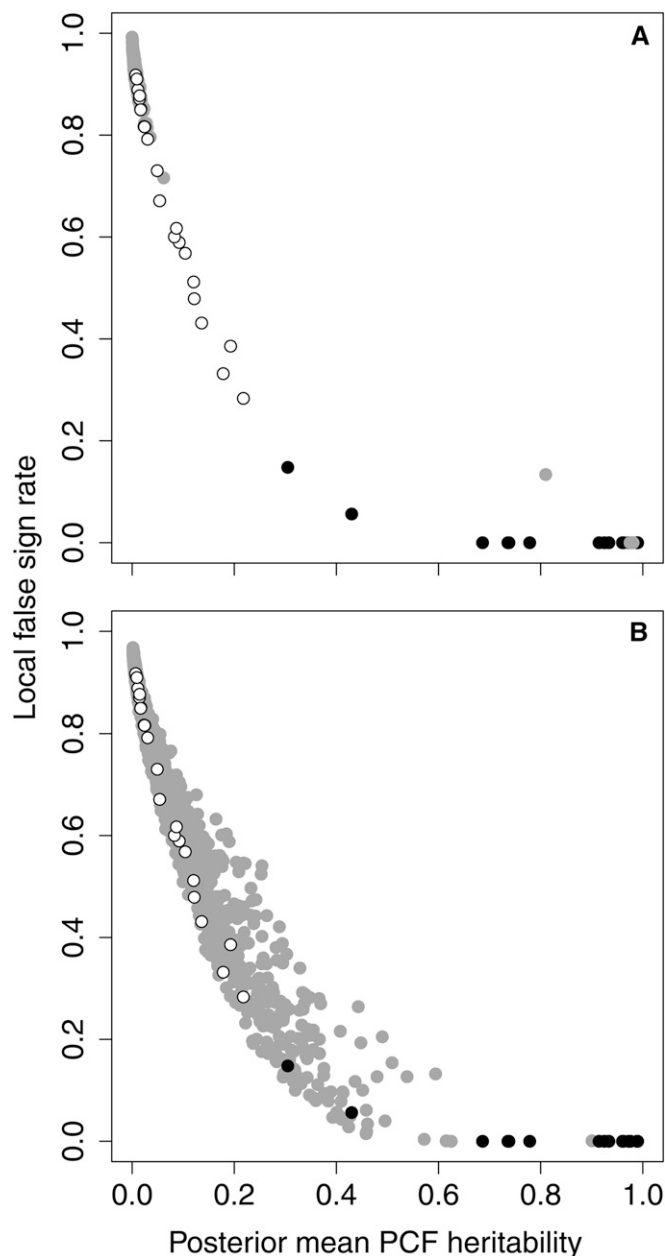
Overall, the 21 significantly heritable PCFs accounted for 46% of the total mutational variance estimated across the 3385 individual expression traits. This number dropped to 45% when the two least heritable PCFs (for which  $LFSR > 0.05$ ) were excluded from the calculation. The remaining mutational variance was accounted for by the 24 nonheritable PCFs (12%) and the 3385 trait-specific mutational variances (42%). Nonheritable PCFs might truly capture mutational covariance, but below the level statistically detectable with our design. Statistically, we cannot interpret the specific mutational variances, as there are 3385 of these parameters and only 40 d.f. at the among-line level. Biologically, the among-line variance partitioned into the specific mutational variance may arise from truly nonpleiotropic mutations and/or the mixture of many pleiotropic mutations of relatively small effect.

To calibrate the biological consequence of the uneven distribution of mutational variance generated by the PCFs,



**Figure 2** Phenotypic common factor (PCF) scores of mutation accumulation line replicate pairs for the 21 heritable PCFs. Scores were estimated within the Bayesian sparse factor model for the two replicate assays from each of the 41 mutation accumulation lines. The solid line shows the 1:1 value. Any deviation of the points from this line reflects the presence of nongenetic (other biological or technical) variation. PCFs are numbered according to their descending phenotypic variance (as per Figure 1).

we adopted a random skewers approach using the BE. The BE is effectively a rotation of the vector of selection through  $\mathbf{G}$ , resulting in the potential for the response to selection to be biased away from the direction of selection (Walsh and Blows



**Figure 3** Comparison of significant and nonsignificant estimated phenotypic common factor (PCF) heritabilities from observed data (black filled and open circles, respectively) with PCF heritabilities from randomized data (gray circles). (A) Type 1 randomizations (same mutational variance as observed data, but null phenotypic and therefore mutational covariance). A single PCF heritability from three of these 100 randomized data sets was in the range of the observed PCF heritabilities. Two of these estimates sit under the single gray point in the bottom right of (A). The remaining 3899 PCF heritabilities from across the 100 data sets were all  $< 0.1$ . (B) Type 2 randomizations (phenotypic covariances as per observed data, but null mutational (co)variance). Of the 5082 estimated PCF heritabilities across the 100 data sets, only 73 (1.4%) were greater than the smallest significant observed PCF heritability.

2009). As the distribution of variance becomes more uneven in  $\mathbf{G}$ , this bias will increase, on average, for random selection gradients. The distribution of angles between a large number of random selection gradients and their corresponding

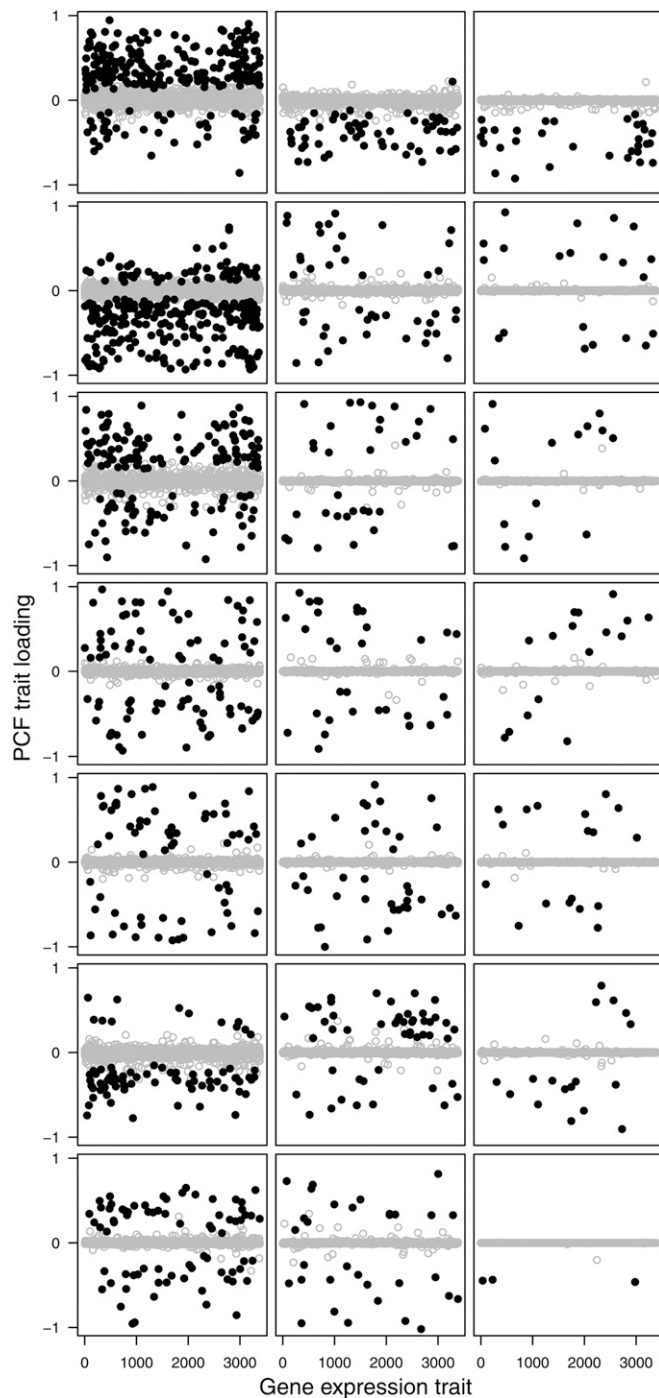
responses therefore allows a quantification of the potential biological impact of the observed distribution of variance. Here, we replaced  $\mathbf{G}$  with  $\mathbf{M}$ , so that we could use the rotation in the BE to determine how the observed distribution of mutational variance might influence evolutionary trajectories of this population.

Figure 5 compares the distribution of angles between the selection gradients and their resulting predicted change in trait means for observed and type 1 randomized data. For both the observed and randomized data, rotating any given random selection gradient through  $\mathbf{M}$  resulted in at least some bias of the response away from the direction of selection. This happens in the randomized data, even though most mutational covariances are small, because the mutational variances are not equal across traits. That is, we expect some nonzero angle between the selection gradient and its response for any  $\mathbf{M}$  that is not exactly proportional to the identity matrix. The majority of angles for the randomized data fell between  $30^\circ$  and  $35^\circ$ , indicating a small to moderate bias on the response to selection. In contrast, the bias in the predicted response to selection away from the direction of selection was consistently  $>76^\circ$  for the observed  $\mathbf{M}$ . Overall, these results suggest that if evolution depended solely on new mutations, uneven magnitudes of mutational variance among individual traits may have modest effects on the direction of phenotypic evolution. In contrast, the observed uneven distribution of mutational variance generated by mutational covariance could strongly bias evolutionary responses if it is in turn reflected in the distribution of standing genetic variance.

Several previous experiments have all reported a similar pattern to that observed here, where mutation does not equally generate variance in all directions of phenotypic space (Camara and Pigliucci 1999; Houle and Fierst 2013; Latimer *et al.* 2014). A caveat of both the current study and those previous investigations is that the distribution of mutational variances uncovered represents only the effect of those mutations that occurred during the generations of experimental MA. It remains to be determined whether multivariate phenotypes with low mutational variance detected in these experiments represent inherent bias in the production of variation or merely stochastic sampling of mutations. Houle and Fierst (2013) compared mutations affecting 20 morphological traits in two independent MA populations. Although trait covariances differed between the two populations, the general pattern that variance was predominantly restricted to relatively few dimensions was common to both, and interestingly, there was evidence that the region of phenotypic space with low mutational variance was shared between the populations (Houle and Fierst 2013), suggesting a repeatable difference in the magnitude of mutational variance across the phenotypic space.

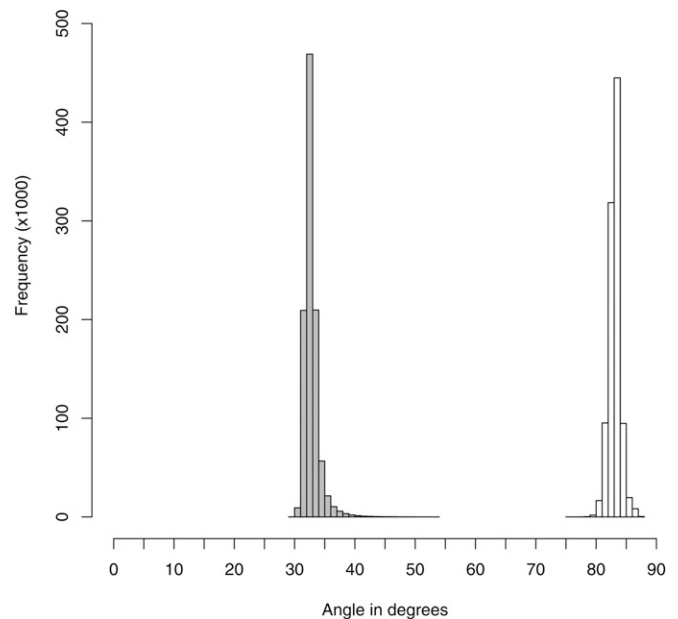
It will be important to determine whether subspaces of low-standing genetic variance can be predicted by subspaces with low mutational variance. The  $\mathbf{G}$ -matrices of natural populations represent mutation–selection–drift balance over longer





**Figure 4** Phenotypic common factor (PCF) trait loadings for the 21 heritable PCFs. PCF trait loadings with  $s < 0.005$  are plotted with black circles. Nonsignificant trait loadings are plotted with gray circles. PCFs are presented in the same order as in Figure 2. Note that PCF trait loadings are limited to an approximate range of  $-1$  to  $1$ , due to the scaling of the data to unit variance.

timescales than MA experiments, where regions of phenotypic space with low genetic variance must represent multivariate phenotypes with either low fitness or little mutational input (Blows and McGuigan 2015). Recently, Houle *et al.*



**Figure 5** Angles between random selection gradients and the corresponding evolutionary response predicted by **M**. The evolutionary responses predicted by the posterior samples of the observed data (open) and the randomized data sets (gray) are each plotted.

(2017) analyzed several *Drosophila* data sets to compare patterns of among-species divergence in wing shape with within species patterns of standing genetic and mutational variance. Both the directions of divergence among species and the distribution of standing genetic variance within a species were strongly explained by the patterns of mutational variance, with directions of maximum divergence and of maximum standing genetic variance corresponding to maximum mutational variance. These results are consistent with there being long-term consequences of the concentration of mutational variance, and might also be interpreted as revealing the evolution of **M** to match the fitness surface (Draghi and Wagner 2008; Pavlicev *et al.* 2011; Jones *et al.* 2014).

A second caveat of the current and previous experiments is that among-line covariances in MA experiments can be generated via both linkage and pleiotropy. Nonpleiotropic covariance can arise when some lines harbor more mutations than others (Keightley *et al.* 2000) and through sampling when the direction of mutation is biased (for example, for life history traits) (Keightley *et al.* 2000; Estes *et al.* 2005). Notably, investigations of nucleotide variation at specific loci indicate the general potential for pleiotropy to generate covariance among hundreds of expression traits (e.g., Gerstung *et al.* 2015; Utrilla *et al.* 2016), as observed here. We further note that the number of traits inferred to be mutationally correlated by our approach is similar to the number of traits inferred by other approaches to be pleiotropically affected by mutation. Recent estimates from several QTL and gene knockout/knockdown studies place the median degree of



pleiotropy in the range 1–9% for morphological and physiological traits, and  $\approx 0.2\%$  for gene expression (Wagner and Zhang 2011). Here, the proportion of traits loading significantly ( $s < 0.005$ ) onto one of the 21 heritable PCFs ranged from  $< 0.1$  to 7.9% with median 1.1% (Figure 4 and Table S3). Nonetheless, further investigation would be required to determine whether the observed concentration of mutational variance was generated through pleiotropic mutation or linkage among individual mutations, each with independent effects on different traits.

Although our results indicate a substantial concentration of the mutational variance in relatively few phenotypic dimensions, the combined effects of the uneven distribution of the mutational variance and their selective effects may result in an even more limited effective dimensionality. Under the framework of Fisher's geometric model, and assuming universal pleiotropy, the empirical distribution of fitness effects of new mutations can be used to estimate  $n_e$ , the effective phenotypic complexity (Martin and Lenormand 2006; Le Nagard and Tenaillon 2013) as:

$$n_e = \frac{p}{1 + CV(\lambda)^2} = 2 \frac{E(s)^2}{V(s)} \quad (4)$$

where  $CV(\lambda)$  is the coefficient of variation of the eigenvalues of  $\mathbf{SM}$ , where  $\mathbf{S}$  is the matrix of selective effects of the mutations on each trait, and  $E(s)$  and  $V(s)$  are the mean and variance of the distribution of the effects of mutations on fitness, respectively. When the eigenvalues of  $\mathbf{SM}$  are all equal, then  $CV(\lambda) = 0$  and all traits are mutationally and selectively independent ( $n_e = p$ ). As  $CV(\lambda)$  increases, the number of independent traits decreases. In this approach, the eigenvalues of  $\mathbf{S}$  and  $\mathbf{M}$  remain unobserved, and  $n_e$  is estimated from empirical estimates of  $E(s)$  and  $V(s)$ . Estimates of  $n_e$  from the distribution of fitness effects are remarkably small, in the range  $0.2 < n_e < 3.0$  for organisms as diverse as viruses and *Drosophila* (Lourenço *et al.* 2011; Tenaillon 2014). However, these are likely to be underestimates if, contrary to a basic assumption of the method, each mutation does not affect all traits (*i.e.*, if pleiotropy is not strictly universal) (Lourenço *et al.* 2011). Direct estimation of  $\mathbf{M}$  and  $\mathbf{S}$  is likely to be required to determine the relationship between  $n_e$  and  $n$ .

The concentration of mutational variance in a small subspace has implications beyond our understanding of evolution in natural populations. Pleiotropic effects of mutations have been demonstrated across different human diseases (Sivakumaran *et al.* 2011; Cross-Disorder Group of the Psychiatric Genomics Consortium *et al.* 2013; Chesmore *et al.* 2018), and there is a growing recognition that taking pleiotropy into account can greatly aid the detection of causal variants and the underlying mechanisms of human diseases (Denny *et al.* 2010; Andreassen *et al.* 2013; Pendergrass *et al.* 2013; Liley and Wallace 2015; Mitteroecker *et al.* 2016). Since we can expect  $n \ll p$  in human populations, genetic analyses that simultaneously consider many widely

disparate disease phenotypes have the potential to greatly enhance our understanding of the impact of mutation on human health.

## Acknowledgments

We thank J. Thompson, C. Oesch-Lawson, D. Petfield, O. Sissa-Zubiate, and Y. H. Ye for help with the mutation accumulation experiment and data collection; A. Lund for help with LaTeX and navigating the Linux environment; and S. Crawley for assistance using the Qriscloud facility. This research was undertaken with the assistance of resources from the Queensland Cyber Infrastructure Foundation (<http://www.qcif.edu.au>). This work was supported by the Australian Research Council.

## Literature Cited

- Andreassen, O. A., W. K. Thompson, A. J. Schork, S. Ripke, M. Mattingdal *et al.*, 2013 Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. *PLoS Genet.* 9: e1003455 [corrigenda: *PLoS Genet.* 11: e1005544 (2015)]. <https://doi.org/10.1371/journal.pgen.1003455>
- Barton, N. H., 1990 Pleiotropic models of quantitative variation. *Genetics* 124: 773–782.
- Blows, M. W., and K. McGuigan, 2015 The distribution of genetic variance across phenotypic space and the response to selection. *Mol. Ecol.* 24: 2056–2072. <https://doi.org/10.1111/mec.13023>
- Blows, M. W., S. L. Allen, J. M. Collet, S. F. Chenoweth, and K. McGuigan, 2015 The phenome-wide distribution of genetic variance. *Am. Nat.* 186: 15–30. <https://doi.org/10.1086/681645>
- Camara, M. D., and M. Pigliucci, 1999 Mutational contributions to genetic variance-covariance matrices: an experimental approach using induced mutations in *Arabidopsis thaliana*. *Evolution* 53: 1692–1703. <https://doi.org/10.1111/j.1558-5646.1999.tb04554.x>
- Chesmore, K., J. Bartlett, and S. M. Williams, 2018 The ubiquity of pleiotropy in human disease. *Hum. Genet.* 137: 39–44. <https://doi.org/10.1007/s00439-017-1854-z>
- Collet, J. M., K. McGuigan, S. L. Allen, S. F. Chenoweth, and M. W. Blows, 2018 Mutational pleiotropy and the strength of stabilizing selection within and between functional modules of gene expression. *Genetics* 208: 1601–1616. <https://doi.org/10.1534/genetics.118.300776>
- Cross-Disorder Group of the Psychiatric Genomics Consortium, Lee, S. H., S. Ripke, B. M. Neale, S. V. Faraone *et al.*, 2013 Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* 45: 984–994. <https://doi.org/10.1038/ng.2711>
- Denny, J. C., M. D. Ritchie, M. A. Basford, J. M. Pulley, L. Bastarache *et al.*, 2010 PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics* 26: 1205–1210. <https://doi.org/10.1093/bioinformatics/btq126>
- Draghi, J., and G. R. Wagner, 2008 Evolution of evolvability in a developmental model. *Evolution* 62: 301–315. <https://doi.org/10.1111/j.1558-5646.2007.00303.x>
- Estes, S., B. C. Ajie, M. Lynch, and P. C. Phillip, 2005 Spontaneous mutational correlations for life-history, morphological and behavioral characters in *Caenorhabditis elegans*. *Genetics* 170: 645–653. <https://doi.org/10.1534/genetics.104.040022>

- Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford. <https://doi.org/10.5962/bhl.title.27468>
- Gerstung, M., A. Pellagatti, L. Malcovati, A. Giagounidis, M. G. Porta *et al.*, 2015 Combining gene mutation with gene expression data improves outcome prediction in myelodysplastic syndromes. *Nat. Commun.* 6: 5901. <https://doi.org/10.1038/ncomms6901>
- Ghosh, J., and D. B. Dunson, 2009 Default prior distributions and efficient posterior computation in Bayesian factor analysis. *J. Comput. Graph. Stat.* 18: 306–320. <https://doi.org/10.1198/jcgs.2009.07145>
- Houle, D., 2010 Numbering the hairs on our heads: the shared challenge and promise of phenomics. *Proc. Natl. Acad. Sci. USA* 107: 1793–1799. <https://doi.org/10.1073/pnas.0906195106>
- Houle, D., and J. Fierst, 2013 Properties of spontaneous mutational variance and covariance for wing size and shape in *Drosophila melanogaster*. *Evolution* 67: 1116–1130. <https://doi.org/10.1111/j.1558-5646.2012.01838.x>
- Houle, D., G. H. Bolstad, K. van der Linde, and T. F. Hansen, 2017 Mutation predicts 40 million years of fly wing evolution. *Nature* 548: 447–450. <https://doi.org/10.1038/nature23473>
- Johnson, T., and N. Barton, 2005 Theoretical models of selection and mutation on quantitative traits. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360: 1411–1425. <https://doi.org/10.1098/rstb.2005.1667>
- Jones, A. G., R. Buerger, and S. J. Arnold, 2014 Epistasis and natural selection shape the mutational architecture of complex traits. *Nat. Commun.* 5: 3709. <https://doi.org/10.1038/ncomms4709>
- Keightley, P. D., E. K. Davies, A. D. Peters, and R. G. Shaw, 2000 Properties of ethylmethane sulfonate-induced mutations affecting life-history traits in *Caenorhabditis elegans* and inferences about bivariate distributions of mutation effects. *Genetics* 156: 143–154.
- Latimer, C. A. L., K. McGuigan, R. S. Wilson, M. W. Blows, and S. F. Chenoweth, 2014 The contribution of spontaneous mutations to thermal sensitivity curve variation in *Drosophila serrata*. *Evolution* 68: 1824–1837. <https://doi.org/10.1111/evo.12392>
- Le Nagard, H., and O. Tenaillon, 2013 Selection-based estimates of complexity unravel some mechanisms and selective pressures underlying the evolution of complexity in artificial networks, pp. 41–61 in *Advances in Network Complexity*, Vol. IV, edited by M. Dehmer, A. Mowshowitz, and F. Emmert-Streib. Wiley-VCH Verlag GmbH & Co, Weinheim, Germany. <https://doi.org/10.1002/9783527670468.ch03>
- Liley, J., and C. Wallace, 2015 A pleiotropy-informed Bayesian false discovery rate adapted to a shared control design finds new disease associations from GWAS summary statistics. *PLoS Genet.* 11: e1004926. <https://doi.org/10.1371/journal.pgen.1004926>
- Lourenço, J., N. Galtier, and S. Glémin, 2011 Complexity, pleiotropy, and the fitness effect of mutations. *Evolution* 65: 1559–1571. <https://doi.org/10.1111/j.1558-5646.2011.01237.x>
- Martin, G., and T. Lenormand, 2006 A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution* 60: 893–907. <https://doi.org/10.1111/j.0014-3820.2006.tb01169.x>
- McGuigan, K., J. M. Collet, E. A. McGraw, Y. X. H. Ye, S. L. Allen *et al.*, 2014 The nature and extent of mutational pleiotropy in gene expression of male *Drosophila serrata*. *Genetics* 196: 911–921. <https://doi.org/10.1534/genetics.114.161232>
- Mitteroecker, P., J. M. Cheverud, and M. Pavlicev, 2016 Multivariate analysis of genotype–phenotype association. *Genetics* 202: 1345–1363. <https://doi.org/10.1534/genetics.115.181339>
- Nadakuditi, R. R., and A. Edelman, 2008 Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples. *IEEE Trans. Signal Process.* 56: 2625–2638. <https://doi.org/10.1109/TSP.2008.917356>
- Orr, H. A., 2000 Adaptation and the cost of complexity. *Evolution* 54: 13–20. <https://doi.org/10.1111/j.0014-3820.2000.tb00002.x>
- Pavlicev, M., J. M. Cheverud, and G. P. Wagner, 2011 Evolution of adaptive phenotypic variation patterns by direct selection for evolvability. *Proc. Biol. Sci.* 278: 1903–1912. <https://doi.org/10.1098/rspb.2010.2113>
- Pendergrass, S. A., K. Brown-Gentry, S. Dudek, A. Frase, E. S. Torstenson *et al.*, 2013 Phenome-wide association study (PheWAS) for detection of pleiotropy within the population architecture using genomics and epidemiology (PAGE) network. *PLoS Genet.* 9: e1003087. <https://doi.org/10.1371/journal.pgen.1003087>
- Runcie, D. E., and S. Mukherjee, 2013 Dissecting high-dimensional phenotypes with Bayesian sparse factor analysis of genetic covariance matrices. *Genetics* 194: 753–767. <https://doi.org/10.1534/genetics.113.151217>
- Sivakumaran, S., F. Agakov, E. Theodoratou, J. G. Prendergast, L. Zgaga *et al.*, 2011 Abundant pleiotropy in human complex diseases and traits. *Am. J. Hum. Genet.* 89: 607–618. <https://doi.org/10.1016/j.ajhg.2011.10.004>
- Stephens, M., 2017 False discovery rates: a new deal. *Biostatistics* 18: 275–294. <https://doi.org/10.1093/biostatistics/kxw041>
- Storey, J. D., 2003 The positive false discovery rate: a Bayesian interpretation and the q-value. *Ann. Stat.* 31: 2013–2035. <https://doi.org/10.1214/aos/1074290335>
- Tenaillon, O., 2014 The utility of Fisher's geometric model in evolutionary genetics. *Annu. Rev. Ecol. Evol. Syst.* 45: 179–201. <https://doi.org/10.1146/annurev-ecolsys-120213-091846>
- Utrilla, J., E. J. O'Brien, K. Chen, D. McCloskey, J. Cheung *et al.*, 2016 Global rebalancing of cellular resources by pleiotropic point mutations illustrates a multi-scale mechanism of adaptive evolution. *Cell Syst.* 2: 260–271. <https://doi.org/10.1016/j.cels.2016.04.003>
- Wagner, G. P., and J. Z. Zhang, 2011 The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nat. Rev. Genet.* 12: 204–213. <https://doi.org/10.1038/nrg2949>
- Walsh, B., and M. W. Blows, 2009 Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. *Annu. Rev. Ecol. Evol. Syst.* 40: 41–59. <https://doi.org/10.1146/annurev.ecolsys.110308.120232>

Communicating editor: E. Stone